

# VIP2024-VAA



Group Project Update 8/27/2024

# Context

- Last week, we were introduced to Prof. Lindshields' work in Senegal on animal behaviour. We delved briefly into the concept of keypoint detection and pose estimation and narrowed our focus to quadruped ungulates.
- The work assigned for the week was to set up GitHub, Slack, and DeadCat. Additionally we had to understand and summarizing given research papers
- Medha and Parth read the AP10K paper, Armaan and Josh read the Open Pose paper, and Claire and Shaan read the Animal Pose paper

# AP10K Dataset Paper

**Title:** *AP-10K: A Benchmark for Animal Pose Estimation in the Wild*

**Authors:** Hang Yu , Yufei Xu , Jing Zhang , Wei Zhao , Ziyu Guan , Dacheng Tao

**Publication year:** 2021

**Publication venue:** NeurIPS Datasets and Benchmarks Track

**What's new in this paper:** The paper contributes one of the first large-scale datasets in animal pose research, containing 10,000 images of animals from 54 species, to allow for increasing generalizability in animal pose estimation. The dataset also contains 50k images without keypoints that are usable for testing.

**Data:** AP-10K

**Method:** The image data is collected through existing datasets, which are then reorganized by taxonomy, and are then manually annotated to include 17 keypoints.

# Open Pose Paper

**Title:** *OpenPose*

**Authors:** Zhe Cao, *Student Member, IEEE*, Gines Hidalgo, *Student Member, IEEE*, Tomas Simon, Shih-En Wei, and Yaser Sheikh

**Publication year:** 2017, last revised 2019

**Publication venue:** CVPR

**What's new in this paper:** The document presents OpenPose, a real-time multi-person 2D pose estimation method that performs bottom-up pose estimation confidence maps and Part Affinity Fields (PAFs). It improves on a previous paper of the same name by refining the PAFs then confidence maps, instead of simultaneously refining each. It details the methodology, multi-person parsing, system overview, and extended foot keypoint detection, and generalization to vehicle pose estimation.

**Data:** MPII, COCO, and COCO+

**Method:** OpenPose uses the VGG-16 model to generate a feature map from the image. It then uses multiple layers of a convolutional neural network to refine part affinity fields for the connections between each keypoint. Then, it takes the refined part affinity fields and generates confidence maps for each keypoint. Finally, it connects the generated keypoints using the PAFs and bipartite matching.

**Link to Relevant Notes:** [https://docs.google.com/document/d/1kBZWf7lplghFu7pBsNWg\\_mPkq58pkv78UGFTzt0SKuM/edit](https://docs.google.com/document/d/1kBZWf7lplghFu7pBsNWg_mPkq58pkv78UGFTzt0SKuM/edit)

# Animal Pose Dataset

**Title:** *Cross-Domain Adaptation for Animal Pose Estimation*

**Authors:** Jinkun Cao, Hongyang Tang, Hao-Shu Fang, Xiaoyong Shen, Cewu Lu, Yu-Wing Tai

**Publication year:** 2019

**Publication venue:** CVPR

**What's new in this paper:** The paper introduces an animal pose dataset and proposes a novel cross-domain adaptation method to improve animal pose estimation. The approach effectively transfers pose knowledge from labeled animal and human datasets to unlabeled animal classes, producing convincing results even with limited data.

**Data:** VOC2011, COCO

**Method:** The paper employs a “Weakly- and Semi- Supervised Cross-domain Adaptation” (WS-CDA) that combines feature extraction, domain discrimination, and keypoint estimation to learn cross-domain shared features. Following this, a “Progressive Pseudo-Label Based Optimization” (PPLO) strategy is used to iteratively improve pose estimation on unlabeled animal data by using high-confidence predictions as pseudo-labels for further training.

# Next Steps

Here are the action items for the following week:

- Login, utilize and comprehend the OpenMMpose library on DeadCat
- Make individual progress report sections on GitHub

# Personal Progress

# Medha

- Joined Slack, GitHub, Box, logged into DeadCat
- Analyzed and summarized AP10K dataset research paper



# Parth

- Joined the Github, Slack.
- Logged into DeadCat and changed password.
- Read the AP10k dataset paper, proposing the AP-10k dataset

# Armaan

- Joined GitHub, Box and Slack
- Created DeadCat account
- Comprehended, interpreted and abstracted the paper on OpenPose
- Researched and understood Bipartite matching which is utilized in PAF kp connections

# Josh Mansky

- Logged into and setup Slack, Github, and the server
- Read and took notes on the OpenPose paper
  - Read previous OpenPose paper for context on part affinity fields and confidence maps
  - Read about non-max suppression used to filter out the non-primary confidence map in OpenPose
  - Read about Bipartite Matching which is used to properly connect the graph of key points for each person in the image
  - Read about VGG-16 to learn the basics of it and its architecture
  - Read about feature maps and the use of max pooling to make the features in the feature maps more defined

# Claire

- Did basic set-up (joined Slack, Box, and GitHub; logged into DeadCat)
- Read and summarized the *Cross-Domain Adaptation for Animal Pose Estimation* paper

# Shaan

- Joined Github Org, logged onto DeadCat, joined Slack
- Read Cross Domain Adaptation for Animal Pose Estimation